A COMPARISON OF FIRE PROBABILITY MAPS DERIVED FROM GIS MODELING AND DIRECT SIMULATION TECHNIQUES

Calvin A. Farris, Research Fire Ecologist
Department of Forest Resources
University of Idaho, FWR 204, Moscow, ID, 83844-1133
Phone: (208) 885-2102
E-mail: farr9475@uidaho.edu

Charles Pezeshki, Associate Professor
School of Mechanical and Materials Engineering
Washington State University, PO Box 642920, Pullman, WA.
Phone: (509) 335-4662
E-mail: pezeshki@mme.wsu.edu

Leon F. Neuenschwander, Professor
Department of Forest Resources
University of Idaho, FWR 204, Moscow, ID, 83844-1133
Phone: (208) 885-2102
E-mail: leonn@novell.uidaho.edu

ABSTRACT

The widespread use of Geographic Information Systems (GIS) for modeling potential fire occurrence reflects the rapid growth of GIS technology and its importance to fire management planning. Despite the proliferation of available modeling approaches, relatively little is known about the factors that influence their relative performance for a given area. In this paper we compared fire probability maps derived from four models of increasing complexity: 1) knowledge-based index model, 2) spatially weighted index model, 3) probability density function based on historic fire data, and 4) direct fire simulation using FARSITE. Our results indicate that the degree to which computationally efficient models can serve as a surrogate to more complex approaches depends on the desired level of "accuracy" and the scale of the analysis. Advantages and limitations of each approach are discussed in relation to these results.

Keywords: fire probability, GIS model, fire simulation

INTRODUCTION

Identifying areas that have a high probability of burning is an important component of fire management planning (Chou 1992). The development of spatially explicit GIS models has greatly facilitated this process by allowing managers to map and analyze variables contributing to fire occurrence across large, unique geographic units. Despite the numerous approaches and methodologies which have emerged for predicting fire probability, information regarding their relative "accuracy" for a given area is generally lacking. Understanding the advantages and limitations of different modeling approaches is important to ensure that management objectives can be achieved with minimal computation and cost (Keane and Long 1998).

A fundamental difficulty associated with comparing fire probability models is the lack of a known "reference" map from which to assess "accuracy." Simply put, it is impossible to know the "right" answer for potential fire distribution across a landscape at any given time because of the high amount of variability of all factors influencing wildland fire. Testing individual models against empirical data would require many years of observation (often defeating the utility of the model for managers), and/or historic data sets which are generally lacking or are statistically inadequate. Moreover, historic or future fire occurrences may not represent current patterns of fire risk in many cases.

One alternative is to simulate fires under current landscape and weather conditions using spatially explicit process models (Green et al. 1991, Clarke et al. 1994, Coleman and Sullivan 1997). Given a large enough number of simulations across a landscape, areas that are most likely to burn can be quantified and different scenarios can be tested (Clarke et al. 1994). This approach differs from traditional GIS-based models that use thematic overlays of landscape variables to infer where a fire is likely occur. Fire probability maps derived from direct fire simulation techniques can then serve as useful references from which to assess the performance of simpler, more computationally efficient GIS models and provide insight into the various factors that determine their performance.

The purpose of this research was to construct and compare fire probability maps derived from three different GIS models and a direct simulation technique. In this paper we provide a brief overview of selected modeling approaches, describe their construction for a sample watershed, and compare the respective maps.

MODELING APPROACHES

A variety of GIS models have been developed for predicting potential fire occurrence (Chuvieco and Congalton 1989, Chou et al. 1990, Chou et al. 1993, Woods and Gossette 1992, Clarke et al. 1994, Salas and Chuvieco 1994, Chuvieco and Salas 1996, Jain et al. 1996, Maselli et al. 1996, McKelvey and Busse 1996, Gouma and Chronopoulou-Sereli 1998). For this study we chose to compare four modeling approaches along a gradient of increasing complexity: 1) knowledge-based index model, 2) spatially weighted index model, 3) probability density function (pdf) model, and 4) direct simulation. Each model varies in the amount of computation, parameterization time, and relative degree of subjectivity, and thus, provides useful benchmarks for comparison. Fire probability maps were constructed independently for the 1,241 hectare Deadhorse Creek Watershed on the Payette National Forest (PNF), central Idaho, USA. The area contains a diversity of fuels and topography making it suitable for testing different models.

Expert-Knowledge-Based Index Model

The first method used to model fire probability in the study area involved applying a cost function developed from expert knowledge and published data to integrate selected landscape variables. The cost function used was linear, with different weighting coefficients based on determination of the relative effect of each variable. Variables were chosen based on their contribution to fire occurrence and data availability (Table 1). A map was constructed for each independent variable showing the weighted distribution of high and low risk factors across the landscape. A composite ignition score (Ig) and spread risk score (Sp) were then calculated using the following equations (refer to Table 1 for codes):

$$Ig = 3r + 2a + 2e + 2s + sc + rd + sd \quad (1)$$

$$Sg = 4f + 3s + 2a + e + c \quad (2)$$

The overall fire probability (Fp) was calculated as the product of the spread and ignition score:

$$Fp = (Sp)(Ig) \quad (3)$$

Ignition Score	*Spread Score*
ridge position (r)	fuel type (f)
slope (s)	slope (s)
aspect (a)	aspect (a)
elevation (e)	elevation (e)
seral cover (sc)	crown closure (c)
distance to roads (rd)	
structure density (sd)	

Table 1. Landscape variables used in cost function equations for the GIS index model.

The scores were normalized between 0-1 and tabulated on a cell by cell basis across the landscape to produce a final map (Figure 1).

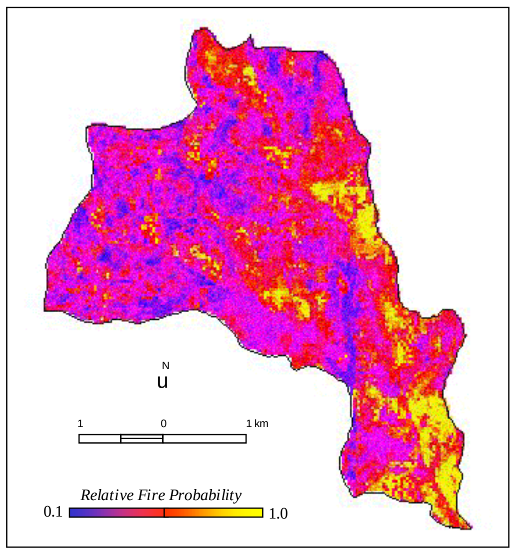

Figure 1. Relative fire probability map produced from the GIS index model.

Basic index models are attractive because they are computationally efficient, many different types and combinations of data can be used, and cost functions can be easily modified to reflect local conditions and expert knowledge. Major drawbacks of index models are that individual cell or polygon values are statistically independent of their neighbor because they are calculated on a unit by unit basis, probability distributions are generally near-normal (rarely observed with empirical fire occurrence data), and cost functions are subject to expert bias.

Spatially Weighted Index Model

Because fire is a contagion process (Chou et al. 1990, Clarke et al. 1994), the likelihood of a given area burning depends not only on its own probability but the probabilities of the areas that surround it (e.g., its spatial neighborhood). Thus, statistical independence associated with index models can be a significant shortcoming when modeling fire probability. One approach to account for this shortcoming is to adjust individual cell probabilities based on the statistical properties of surrounding neighborhoods.

We developed a computationally efficient filtering algorithm to modify the original index model shown in Figure 1. The algorithm calculates the mean probability within successively increasing *n x n* cell neighborhoods in the direction of most likely fire spread (e.g. highest probability). This approximates the statistical distribution of fire spread within an *n* sized neighborhood (or in essence, a fire of that size) (Figure 2a-c). Each neighborhood was assigned a weight corresponding to its relative fire size using fire size distribution curves for the PNF (we used neighborhoods between 1 and 300 ha). The spatially weighted probability (Fpw) for each cell is calculated as follows,

$$Fpw = \sum_{n=1}^{i} (fp_n)(w_n) \quad (4)$$

where fp_n is the cell probability associated with an n sized neighborhood and w_n is the weight assigned to that neighborhood (Figure 2d).

This approach offers the advantage of incorporating spatial dependence to indexed cell probabilities, but it comes with increased computation. It also requires that fire probability values in the original index model be calculated for a large buffer (several km) outside of the target study area to allow for the assessment of adjacent neighborhoods and prevent "boundary" or edge effects. Of course, the overall results are to a large degree dependent on the quality of the original index map.

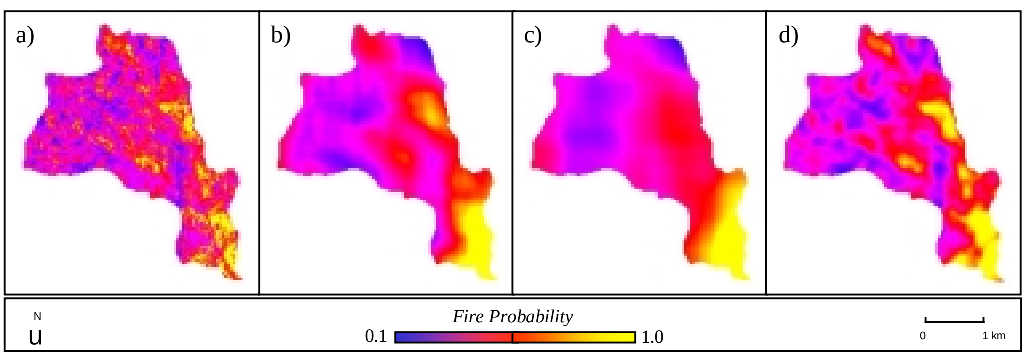

Figure 2. Examples of fire probability maps created during spatial weighting. From left to right are maps for 1 ha (a), 100 ha (b), and 300 ha (c) neighborhoods. Figure 2 (d) is the final spatially weighted map formed through integration of neighborhoods 1-300 ha in size.

Fire Probability Density Function Model

Because of the possibility of expert bias regarding fire frequency and cause, a pdf model was developed using only empirical fire distributions for different landscape variables. In essence, a set of relative probability density functions (pdf) were generated for all the different factors that influence fire (Table 2). Two landscape level joint probability functions were then generated, essentially by multiplying all risk factors for a given geographic area with set characteristics through a table look-up. The first set covered landscape probabilities derived from fire perimeters (n = 439) for the Payette National Forest. The second set covered ignition probabilities (n = 5,693), factoring in all sources of potential ignition, including proximity to roads and human settlement, as well as natural ignitions. These two functions were multiplied to give a combined effects map (Figure 3). By adding all the cumulative relative probabilities across a geographic area and then normalizing by that amount, one ends up with a landscape-level probability distribution.

List of Landscape Variables	
slope	ignition density
aspect	land type
elevation	max temperature
precipitation	land cover
distance to roads	

Table 2. Variables used to calculate empirical ignition and spread probabilities in the pdf model.

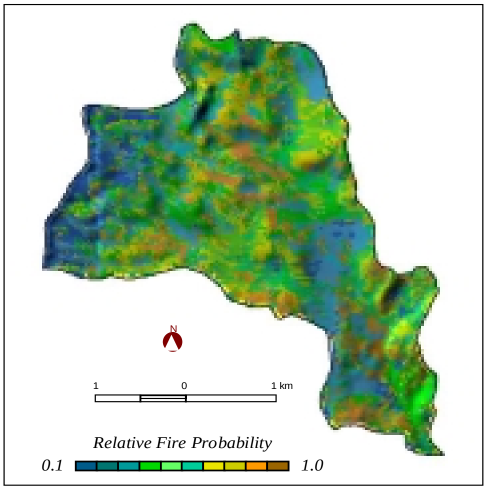

Figure 3. Relative fire probability map derived from the probability density function model.

Although probabilities derived from this approach should give more objective, and presumably more accurate results than the basic knowledge-based index models, there are many limitations. First, spatially explicit fire history data is required. Second, one must assume that historic patterns of fire occurrence represent current fire potential. Third, temporally dynamic variables such as vegetation type or human factors (e.g. roads) can be difficult to assess. Finally, a somewhat arbitrary decision must be made regarding the extent of data to be used when modeling small areas.

Direct Simulation Model

The simplest and most obvious (though temporally and computationally expensive) way of generating relative fire probability maps is to use a fire-perimeter simulator. We used the FARSITE model, developed by Finney (1998) for this study because of its widespread use in the western United States. FARSITE is a wavefront propagation fire extent modeler that models an elliptical diffusion wavefront subject to various inhibiting factors dictated by landscape characteristics. The technique is as follows:

1. Build a model of the landscape using all the various factors required by the FARSITE solver.

2. Model potential site ignition using probabilities from the second half of the landscape-level ignition map generated in the above study.

3. Use Monte Carlo simulation on the above ignition probability density function to start fires across the landscape.

4. Run a statistically significant number of fires out for a specified length of time, tracking their extents and storing this information in a database.

5. Generate a landscape-level histogram that sums the number of times a pixel is "burned" by simulation.

6. Normalize by the total count to generate a landscape level fire probability map (Figure 4).

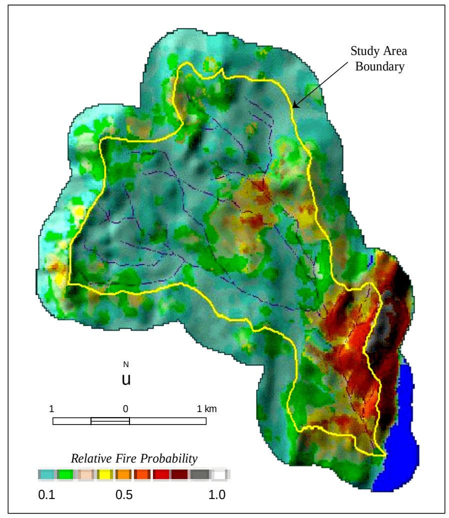

Figure 4. Relative fire probability map derived from the direct simulation model (based on 1,202 simulated fire perimeters using FARSITE).

A total of 1,500 fires were simulated for this experiment, of which 1,202 fires "burned" pixels in the watershed (1,000 originated within the watershed and 202 burned from adjacent ignitions). By allowing adjacent ignitions to burn into the study area, spatial neighborhood influences were incorporated. Simulation locations were generated by flooding the landscape with thousands of potential ignition points. A successful ignition occurred where the predicted ignition probability of a cell exceeded the random probability of a given point. This pattern was supplemented with a systematic ignition grid to ensure a minimum of two ignitions in every 9 ha block. Individual fire duration was assigned randomly from regression curves developed for 1,100 nearby fires on the PNF.

Theoretically, given that the fire extent simulator correctly models fire across the landscape in question, this method gives the "right" or best answer. However, its drawbacks are many. First and foremost are the intensive computational resources required to execute all the fire simulations and the fine-scale weather and fuels data required. Another problem comes with randomizing weather patterns to reflect local conditions. In our simulation runs we assumed worst-case scenarios for weather (mid-August), but only ran the simulations a fixed number of days, in effect, giving a maximal spread but short-lived fire pattern.

MODEL COMPARISON

Two methods were used to compare the respective models. First, Pearson cross correlation coefficients (Zar 1984) were calculated on raw probabilities between the first three models and the direct simulation model. We essentially used this measure as a descriptive index rather than inferential statistic since we had a complete census (e.g., all cells had values). Second, categorical "risk" rankings were compared using error matrices (Jensen 1996), with the direct simulation as the reference image. Quantile rankings were used to convert raw probabilities into four categorical classes (Figure 5). The assignment of names to these categories is somewhat arbitrary given the different underlying statistical distributions of each model. Thus, we only report accuracy for the highest and lowest risk classes, which are generally of most interest to fire managers. Producer's accuracy is the probability that a known class in direct simulation map is correctly classified in the GIS map, and user's accuracy is the probability that a class in the GIS map correctly depicts a corresponding class in the reference (Jensen 1996).

RESULTS

Cross correlation coefficients and categorical "accuracy" results are shown in Table 3. The raw probabilities of all the GIS models were positively correlated with the direct simulation map to some extent, but the spatially weighted index model (r = 0.53) and knowledge-based index model (r = 0.43) had considerably stronger correlation than the pdf model (r = 0.27). A similar overall relationship was observed with the categorical comparisons, although relative differences in

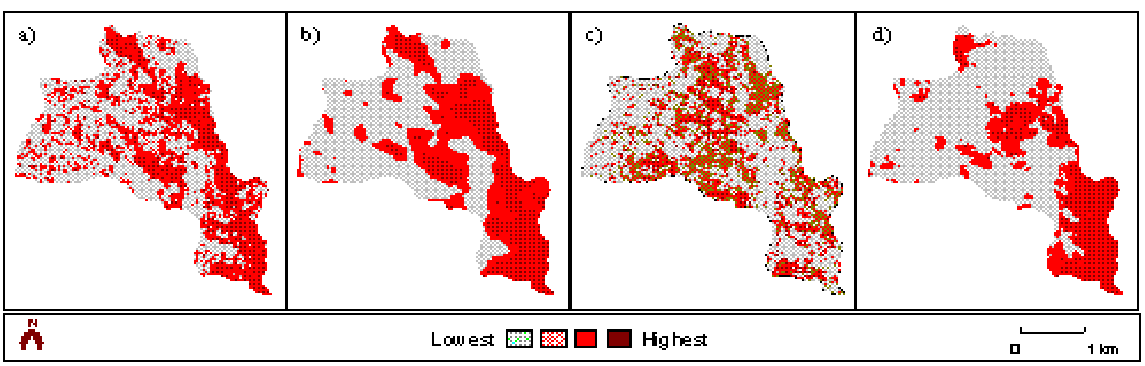

Figure 5. Categorical fire probability maps developed from the basic index model (a), spatially weighted index model (b), probability density function model (c), and direct simulation model (c).

	Correlation	User's Accuracy "high"	User's Accuracy "low"	Producer's Accuracy "high"	Producer's Accuracy "low"
Index Model	0.43	%56	%64	%66	%53
Spatially Weighted Index Model	0.53	%60	%70	%73	%58
Probability Density Function Model	0.27	%34	%45	%31	%37

Table 3. Cross-correlation coefficients (on raw probabilities) and User's and Producer's accuracy for relative risk categories between the three GIS models and the direct simulation model.

accuracy between the knowledge-based index model and spatially weighted model were small. For producer's accuracy, the spatially weighted index model correctly classified 73% of the "high" risk pixels in the simulated map, compared to 66% and 31% for the knowledge-based index model and pdf models respectively. For user's accuracy, approximately 60% of the "high risk" pixels in the spatially weighted index map were classified as "high" risk on the simulated map, compared to 56% and 34% for the knowledge-based index and pdf models. Thus, a relatively large percentage of 'high" risk cells in the direct simulation map were also classified as "high" by the index models. However, the index models slightly over-predicted the total number of "high" risk pixels (Figure 5).

DISCUSSION

The selection of an appropriate model for predicting fire probability ultimately depends on the objectives of the modeling effort. However, managers must carefully weigh computational intensity, complexity, and parameterization requirements against the desired level of “accuracy” (Keane and Long 1998). Such decisions have, in the past, been made difficult by the lack of quantitative data on the relative performance of different models for a given landscape. Often, it is simply assumed that more complex models will produce the most useable results, yet this notion is rarely tested.

In this study, the direct simulation model was clearly the most complex approach in terms of computational intensity and required input parameters (see Finney 1998). Although we believe this technique gives the “best” answer for potential fire distribution in the study area, its utility for broad-scale management applications is limited by its complexity. Moreover, the high level of detail may not be necessary for many management applications. The question then becomes to what degree can less intensive GIS models produce reasonably similar results? Using the direct simulation model as a reference to compare the GIS-based approaches is far from perfect, but in the absence of better data it provides a useful benchmark of computational intensity and state-of-the-art modeling technology.

We did not find a correlation between complexity and “accuracy” among the three GIS models in relation to the direct simulation map. The spatially weighted index model provided the best balance between efficiency and “accuracy” for this particular watershed. Apparently, the basic index model provides a useful initial stratification of high and low probability cell clusters across the landscape, and the filtering expanded upon this pattern to generate more spatially explicit probability distributions based on potential neighborhood influences. As expected, the overall agreement of weighted and unweighted models increases as the level of detail in the output maps is reduced (e.g., converted from discrete to categorical probability). Broad categorical rankings encompass much of the fine scale variability in raw probability distributions between the different models, and are generally more useful for practical fire management planning.

The pdf model produced the most contrasting results relative to other models. One likely explanation is that the empirical probability distributions derived from historic fire perimeters were heavily skewed toward “atypical” fire events. Fewer than ten large fires during two drought years comprised nearly 70% of the total area burned. These fires burned predominantly in cool, mesic, subalpine habitats with little influence from topography (unpublished data). Although such areas burn infrequently, they received a high weight because of the large number of cells represented. Future approaches might involve standardizing scores by fire size or frequency to reduce bias from “weather driven” fires (Bessie and Johnson 1995, Agee 1997). These anomalies were not encountered with the fire ignition data because there was a considerably larger sample size, ignition patterns in the area appear to be more repeatable over time, and each occurrence only counts once (one pixel). For these reasons we believe ignition probabilities are more accurately modeled with this approach, at least in this landscape.

To a large degree, the utility of different modeling approaches will depend on the scale of the analysis. For this study area, fine-scale heterogeneity of simulated fire perimeters contributed significantly to fire probability patterns across the landscape. However, as the average fire size decreases in relation to the size of the study area, the relative importance of this variability will also decrease (for example, variability associated with a 10 ha fire is more important in a 1,000 ha landscape than a 1,000,000 ha landscape). In addition, computational intensity would increase exponentially as the required number of simulations grew. In contrast, larger landscapes are well suited for GIS models because the scale of input data can be adjusted accordingly and increased computation is less dramatic, especially if the objectives are broad (e.g., identify “high” or “low” risk areas).

In summary, no single modeling approach is ideal for all situations and no model will consistently produce accurate results (McKenzie 1998, McKenzie et al. 1996, Keane and Lone 1998). A logical strategy would be to integrate the useful components of various models into hybrid meta-models (Keane et al. 1996). For example, the statistical properties of the simulated fires could be used to calibrate index model cost functions for extrapolation to larger landscapes, or to develop algorithms that impose elliptical neighborhood-based filters in lieu of explicit fire perimeters. Empirical probabilities of individual factors from pdf models can be used for similar purposes (we utilized the ignition component of the pdf model in the direct simulation model and used indexed cell probabilities for spatial weighting). Although these efforts serve to further remove expert bias, qualitative information from expert judgement will always play a role in coarse-scale modeling (McKenzie 1998).

LIMITATIONS

The flexibility that makes GIS models so useful also makes comparisons among them difficult. The models described in this study represent just one possible variant of each general modeling approach. Although slight modifications may have produced different results in some cases, it is virtually impossible to explore all possible combinations of factors and weighting functions. Also, these results were obtained from a single case study in one landscape, so broad generalizations applied to other areas should be made with caution. Using models to validate other models also has many limitations because the assumptions associated with each may compound (McKenzie et al. 1996). Despite these circumstances, we believe that these comparisons can provide useful insight into the various factors that influence different modeling results. These difficulties highlight the challenges involved with comparing and validating models used to predict highly variable natural phenomena such as fire occurrence.

CONCLUSIONS

The degree to which computationally efficient GIS models can serve as a surrogate to more intensive approaches such as direct simulation depends on the level of detail or "accuracy" desired. If the goal is to quantify complex, fine-scale ecological relationships associated with fire spread, then the detailed probability distributions derived from direct simulation techniques would warrant the effort (additional outputs not discussed, such as intensity and rate of spread, might be desirable for other objectives as well). However, if the primary goal is to stratify large landscapes into broad categories (e.g., "high" or "low" risk) to guide fire suppression efforts or prioritize landscape fuels treatments, then more computationally efficient GIS models are a more logical approach. Although the comparisons in this study are far from comprehensive, they provide useful insight into some of the factors that need to be considered when selecting and evaluating a specific model. More research is need in this area.

ACKNOWLEDGMENTS

We thank Liza Fox and Eva Strand of the GIS/Remote Sensing Lab, University of Idaho for their valuable assistance throughout this project. This research was supported in part with funds provided by MacIntire-Stennis, College of Forestry, Wildlife, and Range Sciences, University of Idaho.

REFERENCES

Agee, J. K. (1997). The severe weather wildfire - too hot to handle? *Northwest Science* 71(1):153-156.

Alvarado, E., D. V. Sandberg, and S. G. Pickford. 1998. Modeling large forest fires as extreme events. *Northwest Science*, 72(2):66-75.

Bessie, W. E. and E. A. Johnson. 1995. The relative importance of fuels and weather on fire behavior in subalpine forests. *Ecology,* 76:747-762.

Chou, Y. H. 1992. Management of wildfires with a geographical information system. *International Journal of Geographical Information Systems,* 6(2):123-140.

Chou, Y. H., R. A. Minnich, L. A. Salazar, J. D. Power, and R. J. Dessani. 1990. Spatial autocorrelation of wildfire distribution in the Idyllwild Quadrangle, San Jacinto Mountain, California. *Photogrammetric Engineering & Remote Sensing,* 56(11):1507-1513.

Chou, Y. H, R. A. Minnich, and R. A. Chase. 1993. Mapping probability of fire occurrence in San Jacinto Mountains, California, USA. *Environmental Management*, 17(1):129-140.

Chuvieco, E. and R. G. Congalton. 1989. Application of remote sensing and geographic information systems to forest fire hazard mapping. *Remote Sensing of Environment*, 29:147-159.

Chuvieco, E. and J. Salas. 1996. Mapping the spatial distribution of forest fire danger using GIS. *International Journal of Geographical Information Systems,* 10(3):333-345.

Clarke, K. C.,J. A. Brass, and P. J. Riggan. 1994. A cellular automaton model of wildfire propagation and extinction. *Photogrammetric Engineering & Remote Sensing,* 60(11):1355-1367.

Coleman, J. R. and A. L. Sullivan. 1997. A real-time computer application for the prediction of fire spread across the Australian landscape. *Simulation*, 67(4):230-240.

Finney, M. A. 1998. FARSITE: Fire Area Simulator - model development and evaluation. *RMRS-RP-R.* USDA Forest Service, Rocky Mountain Research Station, Ogden UT.

Gouma, V. and A. Chronopoulou-Sereli. 1998. Wildland fire danger zoning – a methodology. *International Journal of Wildland Fire,* 8(1):37-43.

Green K., M. Finney, J. Campbell, D. Weinstein, and V. Landrum. 1995. Using GIS to predict fire behavior. *Journal of Forestry,* 93(5):21-25.

Jain, A., S. A. Ravan, R. K. Singh, K. K. Das, and P. S. Roy. 1996. Forest fire risk modeling using remote sensing and geographic information system. *Current Science,* 70(10):928-933.

Jensen, J. R. 1996. *Introductory digital image processing: a remote sensing perspective.* Prentice Hall, Upper Saddle River, NJ.

Keane, R. E. and D. G. Long. 1998. A comparison of coarse scale fire effects simulation strategies. *Northwest Science,* 72(2):76-90.

Maselli, F., A. Rodolfi, L. Bottai, and C. Conese. 1996. Evaluation of forest fire risk by analysis of environmental data. *International Journal of Remote Sensing,* 17(7):1417-1423.

McKelvey, K. S. and K. K. Busse. 1996. Twentieth century fire patterns on Forest Service Lands. In *Sierra Nevada Ecosystem Project, Final Report to Congress, v. II, Assessments and Scientific Basis for Management Options*. University of California, Centers for Water and Wildland Resources, Davis, CA.

McKenzie, D., D. L. Peterson, and E. Alvarado. 1996. Extrapolation problems in modeling fire effects at large spatial scales: a review. *International Journal of Wildland Fire,* 6(2):165-176.

McKenzie, D. 1998. Fire, vegetation, and scale: toward optimal models of the Pacific Northwest. *Northwest Science*, 72(2):49-65.

Salas, J. and E. Chuvieco. 1994. Geographic information systems for wildland fire risk mapping. *Wildfire,* 3(2):7-13.

Woods, J. A. and F. Gossette. 1992. A geographic information system for brush fire hazard management. *In Proceedings of the ACSM-ASPRS Annual Convention,* Washington, D. C. pp. 56-65.

Zar, J. H. 1984. *Biostatistical Analysis*. Prentice-Hall International, London (2nd edition), pp. 718.